

## Translations: effects of viewpoint, feature, naming and context on identifying repeatedly copied drawings

Article (Published Version)

Dyson, Benjamin J and Cohen, Rachel (2010) Translations: effects of viewpoint, feature, naming and context on identifying repeatedly copied drawings. *Perception*, 39 (2). pp. 157-172. ISSN 0301-0066

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/52138/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

### **Copyright and reuse:**

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

# Translations: Effects of viewpoint, feature, and naming on identifying repeatedly copied drawings

Benjamin J Dyson

Department of Psychology, Ryerson University, 350 Victoria Street, Ontario, Canada M5B 2K3;  
e-mail: ben.dyson@psych.ryerson.ca

Rachel Cohen

Phoenix Studios, 10–14 Waterloo Place, Brighton BN2 2NB, UK

Received 10 June 2009, in revised form 17 November 2009; published online 10 February 2010

**Abstract.** We explored the tension between bottom–up and top–down contributions to object recognition in a collaboration between a visual artist and a cognitive psychologist. Initial pictorial renderings of objects and animals from various viewpoints were iteratively copied, and a series of drawings that changed from highly concrete images into highly abstract images was produced. In drawing identification in which sets were shown in reverse order, participants were more accurate, more confident, and quicker to correctly identify the evolving image when it was originally displayed from a canonical viewpoint with all salient features present. In drawing identification in which images were shown in random order, more abstract images could be resolved as a result of previously identifying a more concrete iteration of the same drawing. The results raise issues about the influence of viewpoint and feature on the preservation of pictorial images and about the role of labelling in the interpretation of ambiguous stimuli. In addition, the study highlights a procedure in which visual stimuli can degrade without necessitating a substantial loss of complexity.

## 1 Introduction

Our current understanding of visual perception is that it is an active process in which the brain adds to and interprets the often ambiguous retinal information received at the eye (Gregory 1997). This balance between bottom–up and top–down influences on visual processing is particularly apparent during the act of drawing. That is, despite popular techniques for drawing which emphasise the need to simply see what is in front of you without the imposition of semantic labelling or categorisation (Edwards 1999), this principle of the ‘innocent eye’ (Ruskin 1857/1971) has been challenged from a number of directions which emphasise the use of schemata, symbolism, and perceptual history in the production and interpretation of drawing (Gombrich 1977). In the current study we provide an example of collaboration between a visual artist and a cognitive psychologist in which a novel set of iteratively copied drawings was used as stimuli to address the relationship between ambiguity and object recognition, and the implications of active vision on drawing ability.

The association between pictorial representation and object recognition within experimental psychology has been strong, as evidenced by the popularity of various versions of the Snodgrass and Vanderwart picture set [Snodgrass and Vanderwart (1980); see Rossion and Pourtois (2004) for an updated version of the picture set that includes surface detail and colour]. Even with these relatively impoverished two-dimensional (2-D) line drawings, critical manipulations involving viewpoint and labelling have been identified that dramatically influence recognition (see Peissig and Tarr 2007 for a review). It has been well documented that objects are more readily recognised when shown from canonical viewpoints than from non-canonical viewpoints, where the former typically depicts an object from a familiar or typical angle (eg Palmer et al 1981). In addition, identification tends to be quicker for objects when their salient features are present than when they are absent (eg Biederman 1987, figure 22; Hauk et al 2007). While angle typicality and the presence of salient features may be independent contributions

towards the success of object recognition, they have been considered in an integrative manner by some researchers, whereby a canonical view is one that maximises the salience of object features (Bianz et al 1999).

A further influence on the processing of pictorial representations of objects is verbal recoding (see Gombrich 1977, pages 54–64, for an early review). The labelling of objects increases the likelihood of encoding the object in a categorical and view-independent way (Walker et al 2008), which can serve to inhibit future visual analyses (Brandimonte et al 1992; see also Lee 1989; Mitchell et al 2005). The effect of verbal labelling appears to connect more broadly with the psychological phenomenon of verbal overshadowing (Schooler and Engstler-Schooler 1990), referring to the damage caused by the act of verbal description upon subsequent recognition (Lloyd-Jones et al 2006, page 269). Therefore, viewpoint, salient features, and labelling reflect the contribution of numerous influences on the processes involved in object recognition. On the basis of the effects of canonical viewpoint and the presence of salient features, it is clear that we connect what we see with what we expect on the basis of our environmental history. The data on verbal labelling indicate that we attempt to consolidate our current idiosyncratic perceptions into a more schematic form of representation.

While the use of line-drawing stimulus sets is slowly being replaced by more realistic representations of real-world objects and scenes as a result of the increasing availability of digital photography (eg Brady et al 2008; Rousselet et al 2008; St Jacques et al 2008), the use of pictorial representation in understanding the processes associated with object recognition is still important when we consider the act of drawing. In general, our ability to recognise objects in the world far exceeds our ability to pictorially represent them (eg Cohen and Bennett 1997; Kozbelt et al, in press). Nevertheless, we are still able to connect our often schematic scribbles to external portions of our environment even when, as is the case with drawings by young children, the association between drawing and reality is mostly symbolic (Cox 2005, pages 53–55). In this respect, the examination of drawings provides a useful inroad into the processes associated with object recognition, in particular the way in which individuals arrive at decisions about object identification in the face of impoverished or indeterminate pictorial representations.

Early empirical reports into the processing of degraded visual stimuli focused on participants' recognition ability in the face of an increasing number of missing pictorial fragments, such as Gollin's (1960) Incomplete Picture Test (other forms of degradation include visual filtering, eg Vartanian and Goel 2004, and masking, eg Bacon-Macé et al 2005). However, Murray and Szymczyk (1978) suggest that the difficulty associated with object recognition for fragmented images might be less to do with the percentage of image removed and more to do with the absence of distinctive features that coincides with the removal of certain portions of the image (eg Hauk et al 2007). To support this claim, they found that when the degree of fragmentation was controlled, 75% preservation of distinctive features led to better object recognition than 25% preservation of distinctive features. Also of critical importance in this regard is the work of Biederman and colleagues. A specific constraint in identifying degraded objects is whether partial image deletion occurs in a part of the image necessary for the parsing of generalised visual components or geons (Biederman 1987, figure 16). In particular, degraded images become 'non-recoverable' in terms of their identity when the deleted section cannot be reinstated as a function of collinearity or curvature. Such approaches have led researchers towards a perceptual closure account of picture completion (eg Snodgrass and Feenan 1990), in which participants actively attempt to complete the incomplete figure (although see Chikhman et al 2006). This account again describes an entry point for environmental history to influence our raw visual perception, and further demonstrates the reconstructive forces operating during object recognition.

Despite being able to control for the presence or absence of salient features, one potential weakness of the Incomplete Picture Test is the inevitable loss of complexity as the image degrades: in the real world, visual information can be both ambiguous and complex. The current solution to this issue is the adoption of an iterative process by which drawings are copied repeatedly by different participants. Informally referred to as ‘Chinese whispers’, Mesoudi et al (2006) describe this procedure as a critical aspect of cultural transmission and evolution, with iteration producing slight perturbations in information content at each cycle, but cumulatively resulting in large-scale differences.

The novel stimulus set we created here began with representational drawings that were allowed to disintegrate progressively via the act of repeated copying (Bartlett 1932). We propose that, in contrast to the random removal of fragments that guarantees a reduction in visual information, relatively precise iterative copying would potentially control for the upkeep of both stimulus complexity and content. A second and equally important distinction between this and previous work is that researchers have tended to use caricature-like line drawings of objects and animals, in the dominant Snodgrass and Vanderwart (1980) style. Such a pictorial style is likely to make the salient features of the image exaggerated, relative to more artistic renderings employed here that aim to depict objects ‘as seen’ including both salient and subtle visual aspects of the image in question (Murray and Szymczyk 1978). Furthermore, the noise added to the image using the current technique should remain characteristically human in nature. Consequently, after a number of iterations the results could be described as indeterminant in that the images are “detailed and vivid ... that resist object recognition .... highly suggestive of form but not explicitly descriptive of them” (Ishai et al 2007, page 319).

In a primary phase of the experiment, iteration was used as a way to disrupt perceptual information during object recognition by generating increasingly degraded versions of pictorial images of objects and animals initially provided by the artist RC (see figure 1 for an example of the results of the iterative process). In a second phase of the experiment, sets of images were played back in reverse order (cf Snodgrass and Feenan 1990) such that the indeterminate images became increasingly more determinate and eventually returned to the original artist’s rendering, thereby allowing for an assessment of object recognition under progressively improving perceptual conditions. Here we evaluated whether the benefits in object recognition accrued from the presentation of images in canonical perspectives and/or the presence rather than absence of salient features (eg Hauk et al 2007; Palmer et al 1981) also relayed themselves during the act of iterative drawing. Specifically, a canonical, full-featured object after a number of iterations may lead to faster and/or more accurate identification relative to the hypothesised and much more dramatic disintegration of initially non-canonical or feature-absent images. We also examined the generality of the verbal overshadowing effect (Lloyd-Jones et al 2006; Schooler and Engstler-Schooler 1990) by asking participants to either vocalise potential solutions to each degraded image in each series or to vocalise their response only when identification was decided. On the basis of the previous literature, continually naming images should have hindered correct image identification relative to the absence of continual naming. While verbal overshadowing is considered a relatively robust phenomenon applicable to a number of stimuli, including faces and more abstracted figures, the underlying mechanisms for this effect are currently not well understood, although both encoding and retrieval processes have been put forward as contributing factors (see Meissner and Memon 2002 for a brief editorial on the phenomenon). Information regarding the entire stimulus set was collected in two additional conditions where participants responded to each image either in sequential or in random order, in order to assess the potential effects of long-term perceptual priming (Brady et al 2008) between determinant and indeterminant versions of the same object or animal.

## 2 Method

### 2.1 Drawing production

Informed consent for one hundred and thirty-one participants (one hundred and five female) was obtained prior to the experiment and all individuals received £5 per hour or course credit for their time. The mean age within the sample was 21.91 years ( $SD = 4.60$  years) and all reported no formal artistic training. An initial sample of 70 line drawings of familiar objects or animals was produced by the artist (RC) with a black pen on 5 inch  $\times$  3 inch index cards. The only constraint at this stage was that all representations were to be drawn from observation. If a salient feature and/or typical viewpoint could not be confidently identified, the image was drawn once only from a canonical viewpoint and with all features present. Objects and animals deemed viable for manipulation were drawn three times: (a) from a canonical viewpoint, (b) from the same canonical viewpoint but with a salient feature missing, and (c) from a non-canonical viewpoint. Within each drawing session, groups of approximately 8–10 individuals were provided with similar index cards and black pens and asked to copy a random selection of drawings available from the image pool at the time of experimentation. Participants were instructed to copy the drawing as accurately as they could without a time limit. As soon as they had finished a drawing they were given another from a different sequence until an hour was up. This procedure was repeated across groups of participants until each drawing had yielded 19 iterative copies (ie 20 images per drawing; see figure 1).

### 2.2 Drawing identification

**2.2.1 Participants.** Informed consent of sixty participants (thirty-three females) was obtained prior to the experiment and all individuals received £5 per hour for their time. Two participants were replaced in the final sample: one because of inaccurate identification (56% relative to 96% accuracy in the remainder of the group), and one because of sporadic audio-recording failure. Fifteen individuals took part in each of the four experimental conditions, and we also attempted to control for the distribution of age, gender, and artistic training. Mean ages across the four conditions did not significantly vary [ $F_{3,56} = 0.12$ ,  $p = 0.947$ ; average age = 38.15 years ( $SD = 12.15$  years)].

**2.2.2 Stimuli and apparatus.** A subset of 720 drawings (36 sequences of 20) from the drawing production stage was selected for the drawing identification experiment on the basis that a satisfactory level of abstraction was reached. All images were scanned and converted to .pct files with an on-screen size of approximately 4.5 inches  $\times$  2.5 inches. All participants viewed and responded to the images at a comfortable distance from a Macintosh PowerBook G4 computer. Stimulus presentation was controlled with PsyScope (Cohen et al 1993). Of the 36 drawings selected, 27 consisted of the same object or animal drawn three times: at a canonical viewpoint, a canonical viewpoint with a feature missing, and a non-canonical viewpoint (see figure 2 for examples of the three drawing conditions and the results of the iterative copying process). These nine objects or animals (with the missing feature in parentheses) were: bird (beak), boot (eye holes), elephant (trunk), plug (cord), scissors (finger holes), slipper (markings), suitcase (locks), tortoise (shell pattern), and trainer (laces). The remaining nine sets of images drawn from a canonical viewpoint with all features present were: butterfly, chair, hand, jeans, mushrooms, pepper, purse, shirt, and tomatoes.

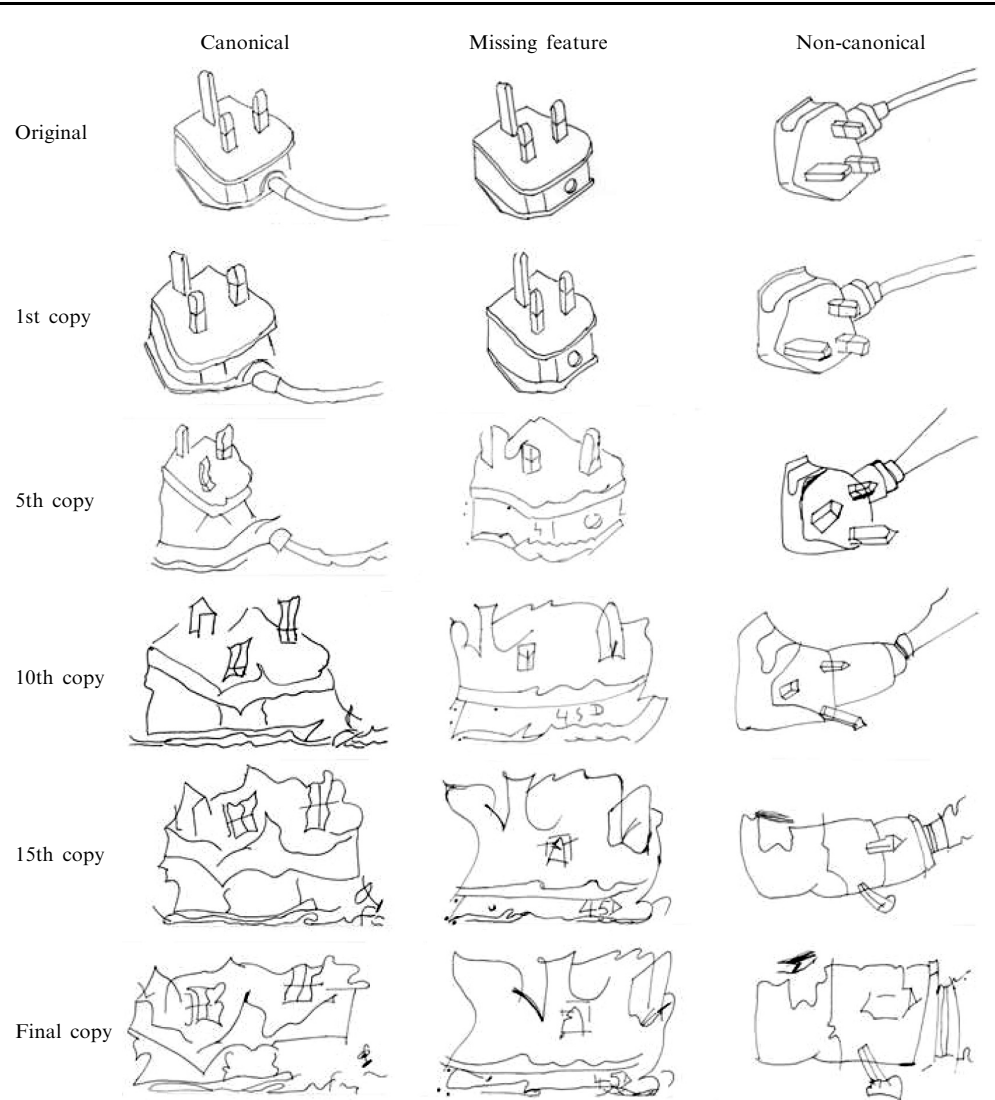
### 2.2.3 Design and procedure

**Set conditions: speaking and silent.** For the set conditions, drawing sets were reversed in order such that participants saw the last iterative copy first and the original drawing last (see figure 1; read bottom to top and right to left). Each drawing was presented



**Figure 1.** Complete example of representation deterioration as a result of 19 iterative copies (reading top to bottom and left to right) from an original pictorial representation of a chair. Note how the mistake in copying on the fourth iteration was not incorporated into the ongoing sequence.

for 3000 ms and immediately replaced by the next image in the sequence. Participants were instructed to press a key as quickly as possible when they thought they knew what the final drawing would represent, with both response time and ratings of confidence (on a 7-point Likert scale; 1 = totally unconfident, 7 = totally confident) being recorded. Verbal responses as to what participants thought the final drawing represented were recorded on Minidisc and later transcribed. After responding, the rest of the drawing set was not shown to the participant (after Snodgrass and Feenan 1990) and the next trial



**Figure 2.** Examples of representation deterioration as a result of iterative copying for canonical, missing-feature, and non-canonical drawings. All initial drawings were copied 19 times to produce drawing sets of 20.

began with the presentation of another drawing set selected at random. Participants first completed a practice block using five drawing sets not included in the experimental series (bananas, cat, flowers, giraffe, and stiletto shoe), followed by an experimental block of the 36 drawing sets described above. During the speaking condition, participants were required to name out loud the possible identity of the current drawing. During the silent condition, participants were not required to name out loud what they thought each current drawing was. Set conditions were completed in a single session lasting approximately 30 min.

**2.2.4 Individual conditions: sequential and random.** For the individual conditions, participants both responded to and named each of the 720 drawings shown in the experimental set conditions. As before, each drawing was presented for 3000 ms during which time participants were encouraged to press a key as quickly as possible when they thought they knew what the drawing represented. Response times, verbal labels,

and confidence ratings were collected as before. If participants failed to respond after 3000 ms, the presentation of the image was terminated and verbal label and confidence ratings were collected as usual. During the sequential condition, participants were shown individual members of the drawing set in reverse order (as in set conditions) while presentation between drawing sets was randomised. During the random condition, the 720 images were presented in a completely randomised order. Individual conditions were completed in a single session lasting approximately 2 h.

3 Results

3.1 Set conditions: speaking and silent

For the dependent variables of image number, feeling of confidence, and accuracy, data were entered into a two-way mixed ANOVA with the within-participants factor of drawing (canonical, non-canonical, missing feature) and between-participants factor of condition (speaking, silent) (see table 1). With respect to the average image number at which participants believed they knew what the final image in the set represented, the ANOVA revealed a significant main effect of drawing ( $F_{2,56} = 104.91, p < 0.001$ ) in the absence of a significant main effect of condition ( $F_{1,28} = 2.61, p = 0.117$ ) and in the absence of a significant interaction between the two ( $F_{2,56} = 0.66, p = 0.521$ ). Tukey's HSD test ( $p < 0.05$ ) revealed that for the main effect of drawing, sets that originated from canonical images led to faster guessing (approximately 12 pictures viewed) than sets that originated from non-canonical (approximately 15 pictures viewed) or missing-feature (approximately 15 pictures viewed) images. In terms of confidence, a main effect of condition was revealed ( $F_{1,28} = 9.39, p = 0.005$ ), in the absence of a main effect of drawing ( $F_{2,56} = 2.87, p = 0.065$ ) and interaction ( $F_{2,56} = 0.32, p = 0.726$ ). Participants in the speaking condition were significantly more confident (6.36) at the time of their response than participants in the silent condition (5.54). Finally, with respect to accuracy, the data yielded no significant effects ( $F_{2,56} = 0.43, p = 0.654$  for the main effect of drawing;  $F_{1,28} = 1.97, p = 0.172$  for the main effect of condition; and  $F_{2,56} = 0.47, p = 0.954$  for the drawing  $\times$  condition term).

**Table 1.** Group mean for average image number, feeling of confidence, and accuracy ratings for the three levels of drawing and two levels of naming condition.

	Image number	Confidence	Accuracy
<i>Continuous naming</i>			
canonical	7.42 (0.54)	6.38 (0.14)	0.96 (0.03)
non-canonical	4.61 (0.38)	6.47 (0.15)	0.97 (0.02)
missing feature	4.35 (0.51)	6.22 (0.20)	0.94 (0.03)
<i>Non-continuous naming</i>			
canonical	8.30 (0.41)	5.52 (0.25)	0.92 (0.03)
non-canonical	5.79 (0.28)	5.62 (0.22)	0.93 (0.03)
missing feature	4.99 (0.45)	5.48 (0.21)	0.91 (0.03)

Note: Images ran in sequences from 20 to 1, hence larger image numbers indicate faster responding. Values in parentheses denote standard error.

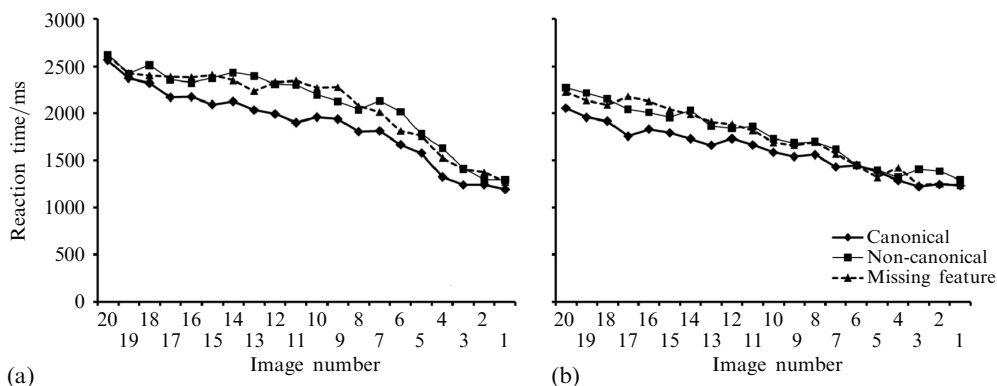
In sum, the data from both set conditions show that, while all participants were accurate in predicting the eventual object or animal, images iteratively copied from canonical, full-feature views yielded an advantage in that participants were able to produce a label consistent with the final image much earlier in the sequence relative to non-canonical or feature-missing sets. The effect of naming produced an additional independent effect in that participants who continually named the images were significantly more confident at the time of responding than participants who did not



explicitly name the images, although this did not impact upon the eventual accuracy of the estimate.<sup>(1)</sup>

### 3.2 Individual conditions: sequential and random

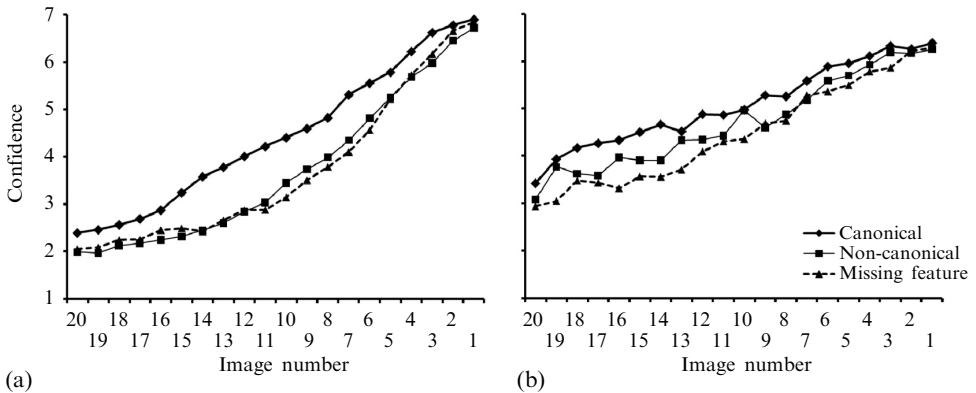
In order to make the data from individual conditions more manageable, responses were collapsed into four bins, each representing the aggregate from five consecutive presentations from indeterminate to determinate images (20–16, 15–11, 10–6, and 5–1, respectively). For the dependent variables of reaction time, feeling of confidence, and accuracy within individual conditions, data were entered into a three-way mixed ANOVA with the within-participants factors of drawing (canonical, non-canonical, missing feature), image bin (20–16, 15–11, 10–6, 5–1), and the between-participants factor of condition (sequential, random). For the reaction-time data (see figure 3), the ANOVA revealed significant main effects for drawing ( $F_{2,56} = 45.66$ ,  $p < 0.001$ ), image bin ( $F_{3,84} = 146.19$ ,  $p < 0.001$ ), and condition ( $F_{2,56} = 4.53$ ,  $p = 0.042$ ), in addition to significant two-way interactions between image bin  $\times$  condition ( $F_{2,56} = 5.03$ ,  $p = 0.003$ ), drawing  $\times$  image bin ( $F_{6,168} = 3.19$ ,  $p = 0.005$ ), and a significant three-way interaction between drawing  $\times$  image bin  $\times$  condition ( $F_{6,168} = 2.98$ ,  $p = 0.009$ ). The only ANOVA terms not to reach statistical significance were the drawing  $\times$  condition interaction ( $F_{2,56} = 1.50$ ,  $p = 0.233$ ). In order to examine this interaction in more detail, effects of drawing and condition across the four image bins were studied with Tukey's HSD test ( $p < 0.05$ ). Broadly speaking, effects of condition were more robust for those images furthest away from the original rendering, with responses in the random condition faster than in the sequential condition. As shown in figure 3, the interaction arose from the observation that faster responses to canonical-view drawings relative to images derived from non-canonical or missing-feature drawings were apparent in earlier image bins for the random condition (20–16, 15–11), while the same reaction-time benefit for canonical-view drawings was apparent in later image bins for the sequential condition (15–11, 10–6, 5–1).



**Figure 3.** Group average reaction times for (a) sequential and (b) random individual conditions across the 20 images within each set. The data show response speed increasing as a result of improved drawing representation.

<sup>(1)</sup> Some concern was raised whether the viewpoints chosen to represent canonical and non-canonical views were considered to be such by a wider audience. Under informal conditions during a gallery show, thirty-eight participants were asked to indicate with respect to each of the nine drawing sets in which canonical and non-canonical viewpoints were available to select the original drawing that best represented the object or animal in question. There was general agreement between the best represented image selected by the participants and the canonical view selected by the artist, although the drawing sets for scissors and suitcase produced some ambiguous results. However, when the data from the set conditions were reanalysed with these drawing sets taken out, the pattern of results was similar, providing modest evidence that the data are not the result of the mislabelling of canonical and non-canonical viewpoints on a subset of images.

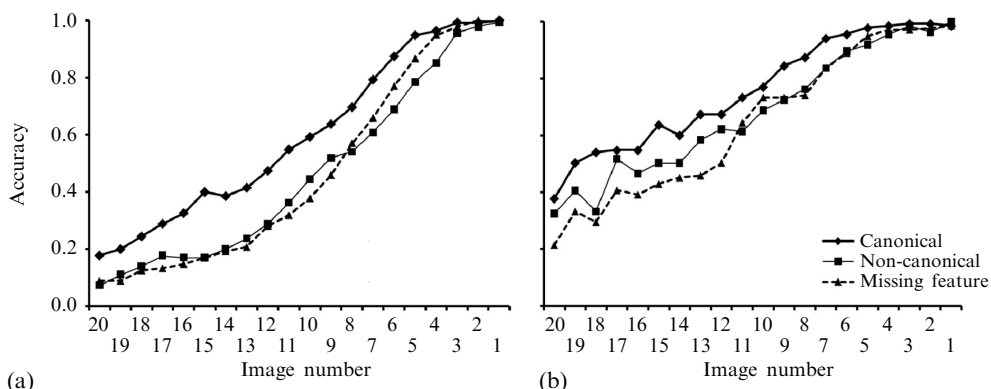
In terms of confidence (see figure 4), all main effects and interactions were significant ( $F_{2,56} = 70.62$ ,  $p < 0.001$  for the drawing main effect;  $F_{3,84} = 339.38$ ,  $p < 0.001$  for image bin main effect;  $F_{1,28} = 4.81$ ,  $p = 0.036$  for the condition main effect;  $F_{6,168} = 9.00$ ,  $p < 0.001$  for the drawing  $\times$  image bin interaction;  $F_{2,56} = 4.93$ ,  $p = 0.011$  for the drawing  $\times$  condition interaction;  $F_{3,84} = 19.68$ ,  $p < 0.001$  for the image bin  $\times$  condition interaction; and  $F_{6,168} = 4.86$ ,  $p < 0.001$  for the drawing  $\times$  image bin  $\times$  condition interaction). Effects of drawing and condition were examined across the four image bins with Tukey's HSD test ( $p < 0.05$ ) in order to reveal the nature of the three-way interaction. In addition to a general increase in confidence as the image got closer to the original, for the earlier image bins (20–16, 15–11) participants in the random condition were generally more confident in their labelling of the image relative to those in the sequential condition, and confidence ratings for images derived from canonical views were generally higher than other drawing types. For the later image bins (10–6, 5–1), confidence ratings between random and sequential conditions were less distinct, and stronger ratings of confidence for canonical images were more apparent in the sequential condition.



**Figure 4.** Group average feeling of confidence regarding the verbal label provided for (a) sequential and (b) random individual conditions across the 20 images within each set. The data show increased confidence as a result of improved drawing representation.

Finally, with respect to accuracy (see figure 5), the ANOVA showed significant main effects of drawing ( $F_{2,56} = 52.60$ ,  $p < 0.001$ ), image bin ( $F_{3,84} = 877.64$ ,  $p < 0.001$ ), and condition ( $F_{1,28} = 52.37$ ,  $p < 0.001$ ), in addition to significant interactions between drawing  $\times$  image bin ( $F_{2,56} = 9.47$ ,  $p < 0.001$ ), and image bin  $\times$  condition ( $F_{3,84} = 30.41$ ,  $p < 0.001$ ). Only the drawing  $\times$  condition interaction ( $F_{2,56} = 2.51$ ,  $p = 0.090$ ) and the three-way interaction ( $F_{6,168} = 1.07$ ,  $p = 0.382$ ) failed to reach significance. We used Tukey's HSD test ( $p < 0.05$ ) to examine the interactions in more detail. For the drawing  $\times$  image bin interaction, accuracy generally improved as the to-be-judged iterative image got closer to the original. Although the final image bin (5–1) showed no significant differences between drawing type (0.98, 0.94, and 0.96 for canonical, non-canonical, and missing feature, respectively), across the remainder of the image bins (20–16, 15–11, and 10–6), accuracy was significantly higher for images originating from a canonical image relative to a non-canonical or missing-feature image. The drawing  $\times$  condition interaction revealed that while accuracy ratings in the final image bin were not significantly different, the random condition yielded more accurate responses for images further along the iterative chain relative to the sequential condition.

In sum, the data from the individual conditions replicate the main findings of the set conditions in that images that originated from canonical drawings yielded faster reaction times and levels of confidence with equivalent levels of accuracy, relative to



**Figure 5.** Group average accuracy of the verbal label provided for (a) sequential and (b) random individual conditions across the 20 images within each set. The data show increased accuracy as a result of improved drawing representation.

images that originated from non-canonical drawings or canonical views with missing features. Consequently, there is little evidence of participants sacrificing accuracy for speed in the current data. Two additional effects were also revealed by this analysis. First, processing facilitation related to canonical drawings was apparent even when a substantial amount of iteration had taken place (ie image bins 20–16 and 15–11). Second, reaction times were faster, participants were more confident and accuracy was higher for the earlier image bins (20–16, 15–11) in the random condition than in the sequential condition.

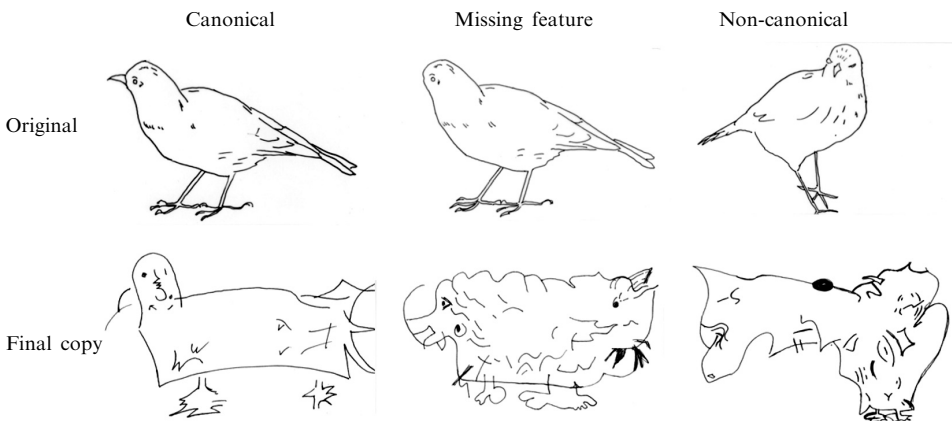
#### 4 Discussion

The results suggest the special nature of drawing sets in which all the salient features of an object or animal were initially presented from a canonical viewpoint. Specifically, participants were able to provide a label consistent with the original rendering much earlier in a series of drawings representing reversed iterative versions of the initial image. While this advantage for canonical, full-feature drawings is consistent with previous research on object recognition (Palmer et al 1981), further experimentation is necessary to clarify whether these drawings were in fact copied more accurately or whether they were simply more recognisable.

It is worth bearing in mind that in the current experiment the participants who copied the drawings were not trained artists. Indeed, the presumption of inaccurate copying (eg Cohen and Bennett 1997) was itself a prerequisite for picture degradation via the iterative process. For many centuries the training of artists in Western cultures has involved drawing from empirical observation, yet there are inherent tensions between so-called bottom-up and top-down contributions to the artistic process (see Kozbelt and Seeley 2007 for a fuller discussion of these issues). In order to accurately represent the world, one approach to artistic training is the development of an ‘innocent’ eye: an attempt to detail what is registered by the eye rather than what is interpreted by the brain (Ruskin 1857/1971). Various strategies are suggested for achieving this, such as copying representational drawings upside down or drawing the spaces between objects rather than the objects themselves (eg Edwards 1999). The emphasis on bottom-up retinal information is consistent with the observation that drawings can become less accurate as the drawer learns more about the object than what is actually seen, such as the tendency for windows and tables to retain a rectangular shape even when seen from oblique views (Cohen and Earls-Jones 2008; Lee 1989; Mitchell et al 2005). However, a second perspective highlights the potential advantages rather than disadvantages of so-called top-down information. As Kozbelt

et al (in press) demonstrate, when asked to create a rendering of a portrait of Samuel Beckett, artists selected local features such as nose, eyes, and mouth whilst non-artists selected more global features such as head and hair outline. The artists' renderings were reliably judged as more accurate, demonstrating that knowledge regarding the selection of key facial features can help to improve the intended representation.

On the basis that top-down contributions might prohibit successful copying in non-artists, a reasonable hypothesis for the current study would have been that objects depicted from non-canonical views or with missing features should have been copied more accurately, being less familiar to the viewer, thereby limiting the role of schemata. Therefore, there are initially inconsistent observations in that (a) better copies should be derived from the lack of top-down influence but (b) canonical full-feature views yielded faster consistent labelling. One explanation for this is that these drawings could have been less accurately copied but that the nature of the inaccuracy was a tendency to redraw familiar features according to schematic (eg Gombrich 1977; Kozbelt et al, in press) pictorial representations. This may have resulted in more caricature-like drawings for the canonical full-feature depictions, yielding consistent labelling at the expense of accurate copying (figure 6 provides one particular example of this from the current series).



**Figure 6.** Comparison of first and last drawings of a bird for canonical, missing-feature, and non-canonical sets.

However, what must also be taken into account is the fact that at particularly advanced stages of iteration, the likelihood of the images retaining particular viewpoints is unlikely. Examination of the most degraded images yields predominantly flat, indeterminant images that are ambiguous both with respect to the representation of viewpoint and feature. One possibility, however, is that what remains of a canonical view after iteration is some form of global perceptual rightness or balance in the image (cf Dyson 2009; Locher 2003). While this accounts for the increased ease of access for consistent labels in the canonical sets relative to non-canonical sets, it currently fails to explain improved performance in canonical full-feature sets relative to canonical missing-feature sets. A clearer metric regarding the accuracy of copying between consecutive images will help to delineate between these competing accounts, and estimates of copying accuracy are currently being explored. Future stimulus sets might also employ only single examples of object and animal categories to afford greater precision in accuracy measures. Consideration must also be given to the idiosyncratic nature of iterative copying, such that if repeated, the same initial artistic renderings would probably not result in the same final abstracted images.

The findings relating to the relationship between verbal and visual image processing were also surprising. The speaking condition was expected to disrupt accurate perceptual

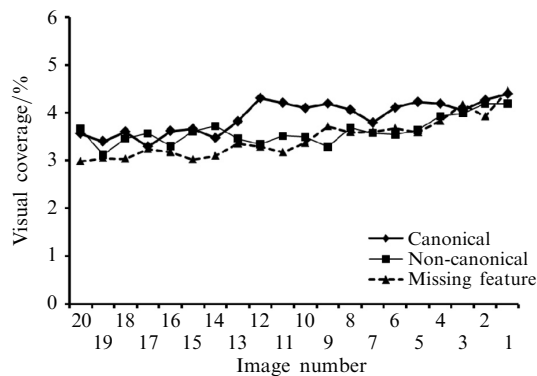
processing (Brandimonte et al 1992; Lee 1989; Mitchell et al 2005), since providing a title for each perceptual interpretation should have increased the likelihood of consolidating and maintaining an incorrect interpretation, even as the object became less ambiguous. However, one possible reason why this effect was absent was that sub-vocalisation was not satisfactorily repressed in the silent condition. In further experiments we would consider a stricter version of articulatory suppression (eg Hanley and Bakopoulou 2003) in place of silence. It has also been suggested that the delirious effects of verbal overshadowing may be limited to facial stimuli when compared with non-faces (Lloyd-Jones et al 2006) and this provides an alternative account of the lack of effect in the present study (although see Meissner and Memon 2002, page 871, for a brief discussion of the generality of the verbal overshadowing effect). Another difference between the current study and previous research is the rate at which verbal responding was requested. In the set conditions, naming was required simultaneously with the presentation of a dynamic collection of images that changed once every 3 s. Consistent with previous research (eg Schooler and Engstler-Schooler 1990), it is possible that reducing the amount of time available for recoding severely attenuated the verbal overshadowing effect. Nevertheless, the observation of higher rating of confidence as a result of continuous naming in the present study does provide an insight into the mechanisms behind overconfidence (Moore and Healy 2008): inviting participants to continually verbalise their current perceptual interpretation seems to be one way of increasing ratings of confidence within an individual.

The current data also reveal that the effects relating to image content were much stronger when participants were required to rate every single picture (ie during individual conditions; sequential and random) relative to responding only when they thought they knew what the final image was (ie during set conditions; speaking and silent). For example, participants were quicker (figure 3), more confident (figure 4), and more accurate (figure 5) in responding to canonical images when presented with individual images in turn, than in making a single response per drawing set in which only global reaction time delineated between viewing conditions (see table 1).

In addition to global differences between the set and individual conditions, a comparison between the two individual conditions also yielded some intriguing outcomes. While responding with equivalent levels of accuracy for image identification, participants in the random condition were both faster and more confident than in the sequential condition. We propose that the locus of this effect stems from the enhanced priming between images in the former condition (Snodgrass and Feenan 1990). Priming was made possible by adopting random image order both within and between sets in the random condition, in that images relatively early on in the initial drawing sets (ie those with a high degree of determinacy) could prime images relatively late on in the initial drawing sets (ie those with a high degree of indeterminacy). Therefore, highly ambiguous images could be resolved as a result of previously identifying an earlier iteration of the same drawing. What is particularly impressive about this type of priming is the maintenance of, and association between, abstract and representational visual information in long-term memory in a condition that lasted approximately two hours (although see Brady et al 2008 for a demonstration of even larger and longer visual long-term memory for objects). In contrast to previously observed and relatively short-lived repetition priming effects for unfamiliar non-verbal stimuli (eg Benton and Moscovitch 1988), the large corpus of stimuli used in the current design (720 images) demanded that any repetition of an image set would take place (on average) across a significant number of trial lags. It is an open question whether the magnitude of priming for these types of stimuli is affected by the iterative distance between the two images, the trial lag between presentations, the indeterminacy of the images, or a combination of all three. Such data are consistent with previous research (eg Biederman

and Cooper 1991, 1992) showing that object priming is relatively insensitive to both size, and translational and reflection viewpoint. In particular, Biederman and Cooper (1992) demonstrated that processing can be facilitated over relatively long temporal lags by a target object that is both a different exemplar and of a different size than the primed object. A more formal analysis of the stimuli produced during the drawing production stage and a systematic manipulation of priming would reveal additional insights into how the visual system manages to find continuity amongst a dynamic and abstract stimulus set.

A final set of issues remains regarding the extent to which iteratively copied drawings preserve stimulus complexity and content, in contrast to the guaranteed reduction in visual information that fragment removal affords (Biederman 1987; Gollin 1960; Murray and Szyczyk 1978). Although there are a number of ways to operationalise visual complexity (eg the number of line segments or line endings), one objective approach is to simply quantify the presence or absence of a mark across each pixel within each image. Consequently, the original .jpg files were submitted to the binary function ( $> 200$ ) using the grey colour map of Matlab (MathWorks), such that the proportion of pixel marks (visual coverage) was calculated for each image. For the dependent variable of visual coverage, data were entered into a two-way between-factors ANOVA with drawing (canonical, non-canonical, missing feature) and image bin (20–16, 15–11, 10–6, 5–1), with each image set (eg bird, suitcase, slipper) representing independent observations. The ANOVA revealed non-significant main effects of drawing ( $F_{2,96} = 0.06, p = 0.941$ ), image bin ( $F_{3,96} = 2.46, p = 0.067$ ), and non-significant interaction ( $F_{6,96} = 0.73, p = 0.623$ ). As can be seen from figure 7, there are slight trends for a reduction in visual coverage as a function of large-scale iteration and also increased visual coverage for mid-range canonical iterations. The data do not indicate significant differences in visual coverage between drawing types for the initially rendered image (image 1 in figure 7), indicating any subsequent and significant modulation of visual coverage to be a function of the interaction between iteration and drawing type rather than anything inherent within the original stimuli themselves. These data offer one way to assess the contribution of visual complexity on iterative image identification, and support the contention that such a process leads to a unique style of degradation in which the semantics of the image is gradually lost but not necessarily at the cost of large amounts of visual information.



**Figure 7.** Estimate of visual coverage of drawing on background as a function of both the degree of iteration (reading right to left: 1 = original image; 20 = 19th iteration) and initial drawing type (canonical, non-canonical, missing feature). The data show trends for a slight reduction in visual coverage as a function of iteration, and slightly greater visual coverage for canonical drawings midway through the sequences.

Further experiments also suggest themselves in terms of testing the stability of content within the image. For example, one could request responding not to image identification but rather to featural (eg respond when the bird loses its beak) or orientation (eg respond when the orientation of the bird is different) changes across images (Biederman 1987). These variations would help to address the relationship between lower-level and higher-level properties of the iterative images. Finally, iterative copying sits alongside more algorithmically controlled filtering, masking, and line perturbation as techniques with which to iteratively degrade images (Bacon-Macé et al 2005; Bartlett 1932; Vartanian and Goel 2004; see also Pearson et al 1990 for a demonstration of computer-generated line drawings). While such manipulations are available, it remains an empirical question whether the results from such algorithmic perturbation would be qualitatively or quantitatively different from the human-driven heuristic perturbation, as shown in figure 1. As stated by Kozbelt et al (in press), it is entirely possible that iterative drawing reveals something about the underlying 'knowledge' of what is being drawn, in a way that cannot be uncovered by using any other manipulation.

In sum, it is argued that, whilst the use of pictorial representation will slowly be surpassed by the use of digital photography as the stimulus of choice in the visual object recognition literature, drawings will continue to provide important expressions of perceptual and cognitive processes. The present study has shown how iterative copying amongst non-artists serves as a novel way to degrade visual stimuli, and not necessarily at the substantial loss of complexity (Murray and Szymczyk 1978), thereby overcoming potential problems with other forms of visual decay such as Gollin's (1960) Incomplete Picture Test. The results showed that while there were very few differences across the various conditions during the first few iterations, these small-scale perturbations in the drawing accuracy of non-artists eventually accumulated in large-scale differences between experimental conditions (Mesoudi et al 2006). In particular, the data reveals that the ability to provide a verbal label consistent with the final representation of an object or animal from essentially abstract and indeterminant images was facilitated by the initial use of an artistic rendering that depicted the object or animal from a canonical, full-feature viewpoint. This study makes it clear that the interaction between bottom-up and top-down processes can be translated from traditional paradigms to the unique domain of drawing, underscoring the necessity to consider the ways in which differences in perception and cognition are embodied and expressed by the individual.

**Acknowledgments.** The research was funded by the British Academy. The authors would like to thank Phoebe Arrowsmith-Brown, Angela Brew, and Michelle Ogborn for additional discussion during the writing of the manuscript, and also Aaron Kozbelt and four anonymous individuals for their comments during the reviewing of the manuscript.

## References

- Bacon-Macé N, Mace M J M, Fabre-Thorpe M, Thorpe S, 2005 "The time course of visual processing: Backward masking and natural scene categorization" *Visual Research* **45** 1459–1469
- Bartlett F C, 1932 *Remembering: A Study in Experimental and Social Psychology* (Cambridge: Cambridge University Press)
- Bentin S, Moscovitch M, 1988 "The time course of repetition effects for words and unfamiliar faces" *Journal of Experimental Psychology: General* **117** 148–160
- Biederman I, 1987 "Recognition-by-components: A theory of human image understanding" *Psychological Review* **94** 115–147
- Biederman I, Cooper E E, 1991 "Evidence for complete translational and reflectional invariance in visual object priming" *Perception* **20** 585–593
- Biederman I, Cooper E E, 1992 "Size invariance in visual object priming" *Journal of Experimental Psychology: Human Perception and Performance* **18** 121–133
- Blanz V, Tarr M J, Bülthoff H H, 1999 "What object attributes determine canonical views?" *Perception* **28** 575–599

- 
- Brady T F, Konkle T, Alvarez G A, Oliva A, 2008 "Visual long-term memory has a massive storage capacity for object details" *Proceedings of the National Academy of Sciences of the USA* **105** 14325–14329
- Brandimonte M A, Hitch G J, Bishop D V M, 1992 "Verbal recoding of visual stimuli impairs mental image transformations" *Memory & Cognition* **20** 449–455
- Chikhman V, Shelepin Y, Foreman N, Merkuljev A, Pronin S, 2006 "Incomplete figure perception and invisible masking" *Perception* **35** 1441–1457
- Cohen D J, Bennett S, 1997 "Why can't most people draw what they see?" *Journal of Experimental Psychology: Human Perception and Performance* **23** 609–621
- Cohen D, Earls-Jones H, 2008 "How shape constancy relates to drawing accuracy" *Psychology of Aesthetics, Creativity, and the Arts* **2** 8–19
- Cohen J, MacWhinney B, Flatt M, Provost J, 1993 "PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers" *Behavior Research Methods, Instruments, & Computers* **25** 257–271
- Cox M, 2005 *The Pictorial World of the Child* (Cambridge: Cambridge University Press)
- Dyson B J, 2009 "Aesthetic appreciation of pictures", in *Encyclopedia of Perception* Ed. B Goldstein (London: Sage)
- Edwards B, 1999 *The New Drawing on the Right Side of the Brain: A Course in Enhancing Creativity and Artistic Confidence* (New York: Tarcher/Penguin)
- Gollin E S, 1960 "Developmental studies of visual recognition of incomplete objects" *Perceptual and Motor Skills* **11** 289–298
- Gombrich E H, 1977 *Art and Illusion: A Study in the Psychology of Pictorial Representation* 5th edition (Oxford: Phaidon Press)
- Gregory R L, 1997 *Eye and Brain: The Psychology of Seeing* 5th edition (Oxford: Oxford University Press)
- Hanley J R, Bakopoulou E, 2003 "Irrelevant speech, articulatory suppression, and phonological similarity: A test of the phonological loop model and the feature model" *Psychonomic Bulletin & Review* **10** 435–444
- Hauk O, Patterson K, Woollams A, Cooper-Pye E, Pulvermuller F, Rogers T T, 2007 "How the camel lost its hump: The impact of object typicality on event-related potential signals in object decision" *Journal of Cognitive Neuroscience* **19** 1338–1353
- Ishai A, Fairhall S L, Pepperell R, 2007 "Perception, memory and aesthetics of indeterminate art" *Brain Research Bulletin* **73** 319–324
- Kozbelt A, Seeley W P, 2007 "Integrating art historical, psychological, and neuroscientific explanations of artists' advantages in drawing and perception" *Psychology of Aesthetics, Creativity, and the Arts* **1** 80–90
- Kozbelt A, Seidel A, El Bassiouny A, Mark Y, Owen D R, in press "Visual selection contributes to artists' advantages in representational drawing" *Psychology of Aesthetics, Creativity, and the Arts*
- Lee M, 1989 "When is an object not an object? The effect of 'meaning' upon the copying of line drawings" *British Journal of Psychology* **80** 15–37
- Lloyd-Jones T J, Brown C, Clarke S, 2006 "Verbal overshadowing of perceptual discrimination" *Psychonomic Bulletin & Review* **13** 269–274
- Locher P J, 2003 "An empirical investigation of the visual rightness theory of picture perception" *Acta Psychologica* **114** 147–164
- Meissner C A, Memon A, 2002 "Verbal overshadowing: A special issue exploring theoretical and applied issues" *Applied Cognitive Psychology* **16** 869–872
- Mesoudi A, Whiten A, Laland K N, 2006 "Towards a unified science of cultural evolution" *Behavioral and Brain Sciences* **29** 329–347
- Mitchell P, Ropar D, Ackroyd K, Rajendran G, 2005 "How perception impacts on drawings" *Journal of Experimental Psychology: Human Perception and Performance* **31** 996–1003
- Moore D A, Healy P J, 2008 "The trouble with overconfidence" *Psychological Review* **115** 502–517
- Murray F S, Szymczyk J M, 1978 "Effects of distinctive features on recognition of incomplete pictures" *Developmental Psychology* **14** 356–362
- Palmer S, Rosch E, Chase P, 1981 "Canonical perspective and the perception of objects", in *Attention and Performance IX* Eds J Long, A Baddeley (Hillsdale, NJ: Lawrence Erlbaum Associates) pp 135–151
- Pearson D, Hanna E, Martinez K, 1990 "Computer generated cartoons", in *Images and Understanding* Eds H Barlow, C Blakemore, M Weston-Smith (Cambridge: Cambridge University Press) pp 46–60



- 
- Peissig J J, Tarr M A, 2007 "Visual object recognition: Do we know more now than we did 20 years ago?" *Annual Review of Psychology* **58** 75–96
- Rossion B, Pourtois G, 2004 "Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic level object recognition" *Perception* **33** 217–236
- Rousselet G A, Husk J S, Bennett P J, Sekuler A B, 2008 "Time course and robustness of ERP object and face differences" *Journal of Vision* **8** 1–18
- Ruskin J, 1857/1971 *The Elements of Drawing* (Mineola, NY: Dover)
- Schooler J W, Engstler-Schooler T Y, 1990 "Verbal overshadowing of visual memories: some things are better left unsaid" *Cognitive Psychology* **22** 36–71
- Snodgrass J G, Feenan K, 1990 "Priming effects in picture fragment completion: Support for the perceptual closure hypothesis" *Journal of Experimental Psychology: General* **119** 276–298
- Snodgrass J G, Vanderwart M, 1980 "A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity" *Journal of Experimental Psychology: Human Learning and Memory* **6** 174–215
- St Jacques P L, Rubin D C, LaBar K, Cabeza R, 2008 "The short and long of it: Neural correlates of temporal order-memory for autobiographical events" *Journal of Cognitive Neuroscience* **20** 1327–1341
- Vartanian O, Goel V, 2004 "Neuroanatomical correlates of aesthetic preference for paintings" *NeuroReport* **15** 893–897
- Walker P, Blake H, Bremner J G, 2008 "Object naming induces viewpoint-independence in longer-term visual remembering: Evidence from a simple object drawing task" *European Journal of Cognitive Psychology* **20** 632–648

ISSN 0301-0066 (print)

ISSN 1468-4233 (electronic)

# PERCEPTION

VOLUME 39 2010

[www.perceptionweb.com](http://www.perceptionweb.com)

**Conditions of use.** This article may be downloaded from the Perception website for personal research by members of subscribing organisations. Authors are entitled to distribute their own article (in printed form or by e-mail) to up to 50 people. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.